

INTERVENTIONAL CARDIOLOGY AND SURGERY

The logistic EuroSCORE in cardiac surgery: how well does it predict operative risk?

F Bhatti, A D Grayson, G Grotte, B M Fabri, J Au, M Jones, B Bridgewater on behalf of the North West Quality Improvement Programme in Cardiac Interventions



Heart 2006;92:1817-1820. doi: 10.1136/hrt.2005.083204

See end of article for authors' affiliations

Correspondence to:
Mr Ben Bridgewater, South
Manchester University
Hospital, Southmoor Road,
Manchester M23 9LT, UK;
ben.bridgewater@smuht.
nwest.nhs.uk

Accepted
10 February 2006
Published Online First
17 March 2006

Objectives: To study the ability of the logistic EuroSCORE to predict operative risk in contemporary cardiac surgery.

Design: Retrospective analysis of prospectively collected data.

Setting: All National Health Service centres undertaking adult cardiac surgery in northwest England.

Patients: All patients undergoing cardiac surgery between April 2002 and March 2004.

Main outcome measures: The predictive ability of the logistic EuroSCORE was assessed by analysing how well it discriminates between patients with differing observed risk by using the area under the receiver operating characteristic (ROC) curve and studying how well it is calibrated against observed in-hospital mortality. The performance of the EuroSCORE was examined in the following surgical subgroups: all cardiac surgery, isolated coronary artery surgery, isolated valve surgery, combined valve and coronary surgery, mitral valve surgery, aortic valve surgery and other surgery.

Results: 9995 patients underwent surgery. The discrimination of the logistic EuroSCORE was good with a ROC curve area of 0.79 for all cardiac surgery (range 0.71-0.79 in the subgroups). For all operations, the predicted mortality was 5.7% and observed mortality was 3.3%. The logistic EuroSCORE overpredicted observed mortality for all subgroups but by differing degrees ($p = 0.02$)

Conclusions: The logistic EuroSCORE is a reasonable overall predictor for contemporary cardiac surgery but overestimates observed mortality. Its accuracy at predicting risk in different surgical subgroups varies. The logistic EuroSCORE should be recalibrated before it is used to gain reassurance about outcomes. Caution should be exercised when using it to compare hospitals or surgeons with a different operative case mix.

Interest in mortality outcomes in hospitals and for patients treated by individual surgeons over recent years has led to publication of named surgeon outcomes for coronary artery surgery and aortic valve replacement in the UK.^{1,2} The case mix is known to vary greatly between both hospitals and surgeons.^{3,4} Some form of adjustment for differences is believed to be important if meaningful comparisons are to be made.

The additive EuroSCORE is a widely used risk prediction algorithm for cardiac surgery.⁵ It allocates incremental risk points to 17 risk factors to give a score that reflects operative mortality. It has a reasonable overall predictive ability for both coronary and valve surgery but is known to overpredict risk compared with observed mortality in contemporary coronary artery surgery. It has also been shown to have poor predictive ability for patients at higher risk and it underpredicts observed mortality for combined valve and coronary surgery.^{3,4,6,7}

More recently the logistic EuroSCORE has been used.^{8,9} This uses a more complex algorithm to derive risk from the same preoperative and operative risk factors. It is claimed to be a better predictor of operative risk than the additive model, but this has not been studied systematically. We have therefore analysed the predictive ability of the logistic EuroSCORE on a large contemporary cardiac surgery database in northwest England.

METHODS

The North West Quality Improvement Programme in Cardiac Interventions is a regional consortium of all four NHS centres (Blackpool Victoria Hospital; The Cardiothoracic Centre,

Liverpool; Manchester Royal Infirmary; South Manchester University Hospital) performing adult cardiac surgery and percutaneous coronary interventions in the north west of England. The goal of the group is to improve continuously the quality of care for patients receiving cardiac interventions by using a regionally based systems approach.¹⁰

Data were collected prospectively on all patients undergoing adult cardiac surgery between 1 April 2002 and 31 March 2004 in the north west of England. Each patient had a dataset collected, which included preoperative and operative variables, to enable a predicted mortality to be calculated. Data were collected in each institution and returned to a central source for analysis. Data were validated in each centre. Mortality was defined as any in-hospital death.

Design of the study

Two specific questions were addressed:

Is the logistic EuroSCORE a good overall predictor of operative mortality for cardiac surgery?

Does the logistic EuroSCORE predict observed mortality equally well in the various surgical subgroups?

Statistical analysis

All data were analysed with SAS for Windows V.8.2 (SAS Institute, Cary, North Carolina, USA). Categorical data are shown as a percentage. Predicted mortality was calculated for each patient by using the logistic EuroSCORE formula.⁹ If a patient factor necessary to calculate the EuroSCORE was missing in the record, that factor was assumed to be absent (occurred in less than 2% of cases).

We have assessed the predictive ability of the logistic EuroSCORE in two ways. Firstly, does the model discriminate appropriately between groups of patients with different observed risk? Secondly, does the numerical value produced by the model accurately reflect observed mortality (that is, is it calibrated appropriately)?

To assess the discriminatory ability of the logistic EuroSCORE we have used the area under the receiver operating characteristic (ROC) curve.¹¹⁻¹² An area of 0.5 reflects no discrimination and an area of 1.0 indicates a perfect predictor. Areas of greater than 0.7 are generally thought to be useful. To check correct calibration of the logistic EuroSCORE we have simply compared predicted with observed mortality.

In addition to studying the ability of the logistic EuroSCORE for all cardiac surgery we have studied its performance in the following surgical subgroups: isolated coronary artery bypass graft (CABG), isolated valve surgery, combined CABG and valve surgery, mitral valve with or without CABG, aortic valve with or without CABG, and other cardiac surgery. The other group includes any operation that is not either a valve or CABG operation or a combination of the two, and includes postinfarct ventricular septal defect repair, major aortic surgery and other miscellaneous operations. The observed to expected mortality ratios in these subgroups were compared by χ^2 test.

Lastly, as one of the criticisms of the additive EuroSCORE is that it does not predict well in higher-risk patients,⁴⁻⁷ we have studied the performance of the logistic EuroSCORE in low and high-risk patients by dividing the groups into two at a cut off of additive EuroSCORE of 6 or more.⁴

RESULTS

In total 9995 patients underwent cardiac surgery during the study period. Table 1 shows the breakdown by operative category. The "other" cardiac surgery group contained 24 ventricular septal defects, 247 major aortic diseases and 472 miscellaneous other operations.

Table 2 shows the predictive ability of the logistic EuroSCORE in the different categories. The area under the ROC curve was satisfactory for all types of cardiac surgery, ranging from 0.71 to 0.79.

The logistic EuroSCORE predicted a mortality that was higher than observed for all operative subgroups. For all cardiac surgery the predicted mortality was 5.7% compared with an observed mortality of 3.3%, meaning that the logistic EuroSCORE needs to be calibrated by a factor of 0.58 to give an accurate representation of operative risk for this population. The calibration required between subgroups varied significantly, ranging from 0.44 for isolated valve surgery to 0.76 for other cardiac surgery ($p = 0.02$).

Table 3 gives the predictive ability of the logistic EuroSCORE in low and high-risk patients. For all types of

Table 1 Breakdown by operative category of 9995 patients undergoing cardiac surgery in northwest England, 1 April 2002 to 31 March 2004

Type of surgery	Number of patients
All	9995
CABG only	6745
Valve only	1523
CABG + valve	984
Aortic valve \pm CABG	1577
Mitral valve \pm CABG	690
Other cardiac surgery	743

CABG, coronary artery bypass graft.

Table 2 Predictive ability of the logistic EuroSCORE by operative subgroup

Type of surgery	Observed mortality	Area under ROC curve	Logistic EuroSCORE	O:E
All	3.3%	0.79	5.7%	0.58
CABG only	2.0%	0.77	3.9%	0.51
Valve only	3.5%	0.79	7.9%	0.44
CABG + valve	7.2%	0.73	9.6%	0.75
Aortic valve \pm CABG	4.4%	0.76	8.0%	0.55
Mitral valve \pm CABG	4.5%	0.76	9.2%	0.49
Other cardiac surgery	9.8%	0.71	12.9%	0.76

CABG, coronary artery bypass graft; O:E, ratio of observed to expected mortality; ROC, receiver operating characteristic.

cardiac surgery the discriminatory ability was equally good in both groups and the logistic EuroSCORE overpredicted observed mortality, but it did so by a similar degree for both low and high-risk patients. The findings were similar for isolated CABG.

DISCUSSION

Principle findings

The logistic EuroSCORE is a reasonable overall predictor for contemporary cardiac surgery but overestimates observed mortality. The model overpredicts observed mortality to different degrees in the various subgroups. Caution should be exercised when comparing mortality directly against the logistic EuroSCORE or when using the model to compare outcomes between hospitals or surgeons with a differing case mix.

Strengths and weaknesses of the study

Scrutiny on measuring mortality outcomes in cardiac surgery of hospitals and surgeons is increasing.¹⁻³ To allow fair comparisons to be made it is important to adjust observed mortality for predicted risk, and this should be done with a valid risk adjustment model. We have assessed the predictive ability of the logistic EuroSCORE on a large contemporary cardiac surgery database, to which four hospitals and 25 surgeons contribute. The data have been validated locally and have the confidence of clinicians but have not been validated externally, which is a weakness of our study. The study sample is large and therefore we believe our findings are robust.

We have shown that the logistic EuroSCORE overpredicts observed mortality for all operative groups. One possible reason for this is that the mortality after surgery in our region is low and that, rather than our findings suggesting poor

Table 3 Predictive ability of the logistic EuroSCORE in patients at low and high risk

Type of surgery	Low risk (≤ 5 additive EuroSCORE)	High risk (> 5 additive EuroSCORE)
All cardiac surgery		
Number of patients	6747	3248
Observed mortality	1.3%	7.5%
Expected mortality	2.3%	12.8%
O:E	0.56	0.58
ROC	0.69	0.72
Isolated CABG		
Number of patients	5481	1264
Observed mortality	1.1%	5.7%
Expected mortality	2.2%	11.1%
O:E	0.5	0.51
ROC	0.71	0.72

CABG, coronary artery bypass graft; O:E, ratio of observed to expected mortality; ROC, receiver operating characteristic.

performance of the risk model, they suggest good performance of the hospitals in northwest England. However, the reported UK mortality for isolated CABG surgery, isolated valve surgery and combined valve and coronary surgery in 2003 is 2.0%, 4.3% and 7.6%, respectively³; these are similar to the rates reported here. Our non-adjusted surgical results are indicative of the national picture, and there is no reason to suppose that we are achieving mortality figures in line with national data while operating on a significantly higher-risk population than elsewhere in the UK.

Strengths and weaknesses of the study compared with other studies

There are several reports on the performance of risk tools in cardiac surgery. The first risk prediction model to find general use was the Parsonnet score.¹³ This was used by hospitals and surgeons to compare risk-adjusted outcomes with a benchmark,¹⁴ but it has now been shown to significantly overpredict observed mortality.^{3, 15} The additive EuroSCORE was first introduced in 1999 and has been studied widely.⁵⁻⁷ It is now clear that it is not accurate at predicting mortality in higher-risk patients and it underestimates observed mortality for combined valve and coronary surgery.^{4, 6, 7}

To respond to these limitations the logistic EuroSCORE was introduced in 2003 and was claimed to be a better predictor of mortality for high-risk patients than the additive model.^{8, 9} It has been shown to be a reasonable predictor of risk (and better than the additive score for high-risk patients) in both US and Italian studies,^{16, 17} but the US study suggested recalibration may be needed. The model has been shown to overpredict observed risk in a single-centre Australian report.¹⁸ The logistic EuroSCORE has also been reported to have limitations in predicting outcome for thoracic aortic surgery and a modified score has been suggested.¹⁹ Our large multicentre study confirms the previous claims that the logistic EuroSCORE has good predictive ability in both low- and high-risk groups of patients. However, we have also shown clearly that the logistic EuroSCORE overpredicts observed mortality in UK practice and we have given useful contemporary calibration figures for the various operative subgroups, which we feel will be useful for benchmarking outcomes.

Meaning of the study

Our findings should give cause for caution when using the logistic EuroSCORE as a risk adjustment model for comparing institutional or individual surgeon outcomes. Our observation that the logistic EuroSCORE overestimates observed mortality for isolated coronary artery surgery by a factor of two means that it would be easy to gain false reassurance by comparing observed mortality with that predicted by the algorithm. In fact, to be better than average, you would have to obtain a mortality that was less than half that predicted by the logistic EuroSCORE.

The finding that the logistic EuroSCORE overpredicts mortality is not surprising. The model was developed on data in the late 1990s and surgical practice has changed since then with increasing numbers of patients with coronary disease treated percutaneously. Advances in surgery, anaesthetic and intensive care have led to overall reductions in surgical mortality, despite higher predicted operative mortality.³ A model that was calibrated accurately in the 1990s is therefore unlikely to remain accurate. The present study gives appropriate contemporary calibration factors for the model for surgery in northwest England.

Initiatives to benchmark surgical outcomes and performance in the UK have so far focused on CABG surgery and to a lesser extent aortic valve replacement.¹⁻³ The proportion of surgeons' practice that is high risk has been shown to vary

greatly between surgeons, and valid comparisons can be made only after adjustment for risk by an appropriate model.⁴ Cardiac surgical practice is becoming increasingly specialised, with some surgeons developing major proportions of their workload, for example, in major aortic surgery or mitral valve repair. The benchmark operation of isolated CABG may be only a relatively small amount of a hospital's or surgeon's workload and to rigorously compare outcomes it seems appropriate to include all cardiac surgery undertaken. Our findings of significant differences in calibration for the logistic EuroSCORE in the different operative groups means that the model is not suitable for such purposes without significant modifications. However, the calibration factors we have reported may prove to be useful in allowing hospitals or surgeons to identify risk-adjusted benchmarks against our data.

Unanswered questions and future research

Our study has shown some strengths and weaknesses of the logistic EuroSCORE and should prove useful to clinicians who are interested in studying comparative outcomes. Although we have shown variation in the predictive ability of the model in various subgroups, we have not looked at highly specialised areas such as aortic surgery or mitral repair surgery. The logistic EuroSCORE, once calibrated appropriately, may be of use in these groups or specific risk prediction models may be required for these operations. Although we have shown limitations of the model to adjust for risk and believe this may have implications when comparing outcomes between surgeons or hospitals with a differing case mix, we do not know how important this effect will be, and it would warrant further investigation.

ACKNOWLEDGEMENTS

This study was conducted on behalf of the North West Quality Improvement Programme in Cardiac Interventions. The consultant surgeons involved are John Au, Ben Bridgewater, Colin Campbell, John Carey, John Chalmers, Walid Dhimis, Abdul Deiraniya, Andrew Duncan, Brian Fabri, Elaine Griffiths, Geir Grotte, Ragheb Hasan, Tim Hooper, Mark Jones, Daniel Keenan, Neeraj Mediratta, Russell Millner, Nick Odom, Brian Prendergast, Mark Pullan, Abbas Rashid, Franco Sogliani, Paul Waterworth and Nizar Yonan. The following surgeons left the collaboration during the study period: Narinda Bhatnagar, Albert Fagan, Bob Lawson, Udin Nkere, Peter O'Keefe, Richard Page, Ian Weir and David Sharpe.

We acknowledge the assistance of the audit officers working in each centre for their hard work in collecting and validating the data.

Authors' affiliations

F Bhatti, G Grotte, Manchester Royal Infirmary, Oxford Road, Manchester, UK

A D Grayson, B Fabri, The Cardiothoracic Centre, Thomas Drive, Liverpool, UK

J Au, Blackpool Victoria Hospital, Blackpool, UK

M Jones, B Bridgewater, South Manchester University Hospital, Manchester, UK

Competing interests: BB is a Society of Cardiothoracic Surgeons of GB and Ireland representative on the joint Society of Cardiothoracic Surgeons, Healthcare commission, Department of Health group defining national cardiac surgical audit. BB, AG, GJG, BF and MJ are all members of the steering group of the North West Regional Quality Improvement Programme in Cardiac Interventions.

REFERENCES

- 1 Anon. Guardian Unlimited. Heart surgery data: non risk adjusted, <http://society.guardian.co.uk/nhsperformance/story/0,,1439210,00.html> (accessed 21 Sep 2005).
- 2 Bridgewater B, on behalf of the Adult Cardiac Surgeons of North West England. Mortality data in adult cardiac surgery for named surgeons: retrospective examination of prospectively collected data on coronary artery surgery and aortic valve replacement. *BMJ* 2005;**330**:506-10.

- 3 Keogh BE, Kinsman R. *Fifth national adult cardiac surgical database report 2003*. Reading: Dendrite Clinical Systems, 2004.
- 4 Bridgewater B, Grayson AD, Jackson M, et al. Surgeon specific mortality in adult cardiac surgery: comparison between crude and risk stratified data. *BMJ* 2003;**327**:13-7.
- 5 Roques F, Nashef SA, Michel P, et al. Risk factors and outcome in European cardiac surgery: analysis of the EuroSCORE multinational database of 19,030 patients. *Eur J Cardiothorac Surg* 1999;**15**:816-23.
- 6 Karthik S, Srinivasan AK, Grayson AD, et al. Limitations of additive EuroSCORE for measuring risk stratified mortality in combined coronary and valve surgery. *Eur J Cardiothorac Surg* 2004;**26**:318-22.
- 7 Gogbashian A, Sedraykan A, Treasure T. EuroSCORE: a systematic review of international performance. *Eur J Cardiothorac Surg* 2004;**25**:695-700.
- 8 Roques F, Michel P, Goldstone AR, et al. The logistic EuroSCORE. *Eur Heart J* 2003;**24**:1-2.
- 9 Michel P, Roques F, Nashef SA, et al. Logistic or additive EuroSCORE for high-risk patients? *Eur J Cardiothorac Surg* 2003;**23**:684-7.
- 10 Anon. The Cardiothoracic Centre - Liverpool. North West Quality Improvement Programme in Cardiac Interventions. www.nwheartaudit.nhs.uk (accessed 20 Oct 2005).
- 11 Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982;**143**:29-36.
- 12 Swets JA. Measuring the accuracy of diagnostic systems. *Science* 1988;**240**:1285-93.
- 13 Parsonnet V, Dean D, Bernstein AD. A method of uniform stratification of risk for evaluating the results of surgery in acquired heart disease. *Circulation* 1989;**79**:13-12.
- 14 Nashef SA, Carey F, Silcock MM, et al. Risk stratification for open heart surgery: trial of the Parsonnet system in a British hospital. *BMJ* 1992;**305**:1066-7.
- 15 Wynn-Jones K, Jackson M, Grotte GJ, et al. Limitations of the Parsonnet score for measuring risk stratified mortality in the north west of England. The North West Regional Cardiac Surgery Audit Steering Group. *Heart* 2000;**84**:71-8.
- 16 Zingone B, Pappalardo A, Dreas L. Logistic versus additive EuroSCORE: a comparative assessment of the two models in an independent population sample. *Eur J Cardiothorac Surg* 2004;**26**:1134-40.
- 17 Jin R, Grunkemeier GL. Additive vs. logistic risk models for cardiac surgery mortality. *Eur J Cardiothorac Surg* 2005;**28**:240-3.
- 18 Yap CH, Mohajeri M, Ihle BU, et al. Validation of Euroscore model in an Australian patient population. *ANZ J Surg* 2005;**75**:508-12.
- 19 Barmettler H, Immer FF, Berdat PA, et al. Risk-stratification in thoracic aortic surgery: should the EuroSCORE be modified? *Eur J Cardiothorac Surg* 2004;**25**:691-4.

IMAGES IN CARDIOLOGY

doi: 10.1136/hrt.2006.088260

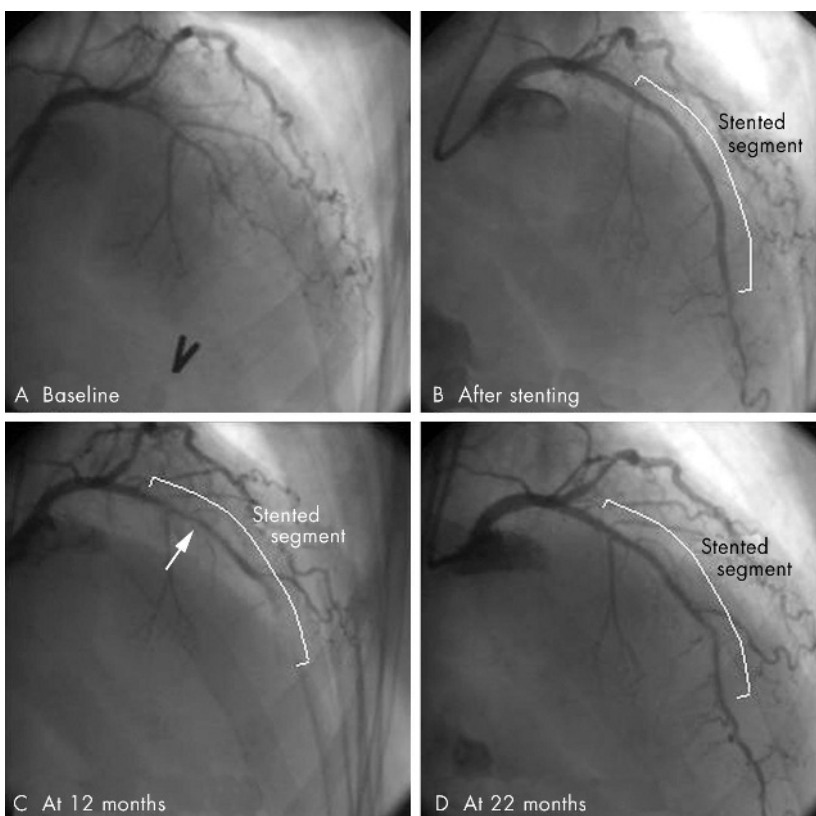
Late thrombosis or in-stent restenosis?

A 55-year-old hypertensive and hypercholesterolaemic woman presenting with non-Q wave myocardial infarction received four 18 mm long sirolimus eluting stents for diffuse atherosclerosis of the left anterior descending (LAD) coronary (panel A, pretreatment; panel B, after stenting). She received double antiplatelet treatment (aspirin and clopidogrel) for six months, after which clopidogrel was stopped.

At one year, routine exercise scintigraphy showed inducible anterior myocardial ischaemia. Three weeks later, angiography indicated total occlusion (type IV restenosis) of the mid LAD (panel C) with type III (diffuse) restenosis of the mid-proximal segment (arrow). The distal LAD received collateral flow from the right coronary artery. Because the patient had been absolutely asymptomatic in the past months, in-stent thrombosis was assumed to be unlikely and it was decided not to attempt repeat revascularisation.

Another 10 months later, the patient was admitted for unstable angina. At this time, angiography showed complete recanalisation of the LAD with residual type II (proliferative) restenosis (panel D), and no collateral circulation from the right coronary artery. Balloon-only angioplasty was performed with a good final outcome.

In-stent thrombosis most commonly presents acutely, with a > 70% incidence of myocardial infarction and > 25% of death. In the case presented here, both the clinical presentation and the angiographic pattern were more consistent with a type III-IV in-stent restenosis, but (late) in-stent thrombosis might have mimicked this angiographic appearance, and the collateral flow might have masked the acute symptoms. We conclude that, because the possibility of in-stent thrombosis cannot be discarded a priori, cardiologists should always consider performing repeat revascularisation, including in the case of totally occlusive lesions.



M Fineschi
T Gori
C Pierli
A Bravi

tommaso.gori@utoronto.ca